ciation with second timing information, wherein the calibration data specifies position information for at least the depth information;

decompressing the received first compressed version of the capture data into a second version of the capture data comprising depth data and audio data;

generating a 3D representation based on the depth data and the calibration data,

wherein one or more 2D images are rendered from the 3D representation at one or more determined viewpoints; and

causing output, by the receiving 3D system, of the audio data and of the one or more 2D images, synchronized based on the first timing information and the second timing information.

9. The computer-readable storage medium of claim **8**, wherein the first compressed version of the capture data was transmitted by a mobile phone device and wherein the receiving 3D system performs the outputting of the 2D images and is an artificial reality device with one or more of:

at least one binocular display;

at least one lightfield display;

at least one holographic display;

at least one wave front display;

at least one 3D stereo displays; or

any combination thereof.

10. The computer-readable storage medium of claim **8**, wherein the operations further comprise:

identifying conversation context factors comprising one or more of:

a determined compute capability of the receiving 3D system;

a determined battery level;

determined display capabilities of the receiving 3D system; or

any combination thereof; and

matching one or more factors of the conversation context to a 3D representation type in a hierarchy of 3D representation types;

wherein the generating the 3D representation comprises generating a 3D representation of the matched 3D representation type.

11. The computer-readable storage medium of claim **8**,

wherein the first compressed version of capture data is associated with calibration data that specifies intrinsic parameters and extrinsic parameters of one or more capture devices that captured the capture data; and

wherein the generating the 3D representation comprises applying one or more transformations to the depth data based on the intrinsic parameters and extrinsic parameters to generate one or more of a point cloud, a signed distance function, populated voxels, a mesh, a light field; or any combination thereof.

12. The computer-readable storage medium of claim **8**, wherein the one or more determined viewpoints are viewpoints of each eye of a user of the receiving 3D system, determined based on a head tracking system and/or an eye tracking system of the receiving 3D system.

13. A sending computing system for providing a 3D conversation, the sending computing system comprising:

one or more capture devices;

one or more processors; and

one or more memories storing instructions that, when executed by the one or more processors, perform operations comprising:

obtaining capture data, from the one or more capture devices, comprising color information, depth information, and audio information;

compressing the capture data into a first version of the capture data; and

providing the first version of the capture data to a reconstruction system that:

receives the first version of the capture data;

decompresses the received first version of the capture data into a second version of the capture data comprising depth data, color data, and audio data; and

generates a 3D representation based on the depth data;

wherein one or more 2D images are rendered from the 3D representation at one or more determined receiving user viewpoints, wherein the one or more 2D images are further based on the color data; and

wherein output of the audio data and display of the rendered one or more 2D images are synchronized.

14. The sending computing system of claim **13**, wherein the one or more capture devices comprise at least a color camera, a depth camera, and a microphone.

15. The sending computing system of claim **13**,

wherein the depth information comprises one or more point clouds; and

wherein compressing the capture data into the first version of the capture data comprises applying a video codec to at least the color information and applying a point cloud compression algorithm to the one or more point clouds.

16. The sending computing system of claim **13**, wherein the operations further comprise:

receiving calibration data associated with particular ones of the one or more capture devices, wherein the calibration data associated with a particular capture device specifies intrinsic parameters and extrinsic parameters of that capture device; and

associating parts of the calibration data with parts of the capture data based on correspondences of a part of the calibration data having been generated for the capture device that captured the associated part of the capture data;

wherein the generating the 3D representation comprises applying one or more transformations to the depth data based on the intrinsic parameters and/or extrinsic parameters.

17. The sending computing system of claim **13**, wherein the reconstruction system is part of a receiving artificial reality system that performs the rendering and the synchronized output of the audio data and display of the one or more 2D images.

18. The sending computing system of claim **13**, wherein the reconstruction system transmits a version of the 3D representation to a receiving artificial reality system that performs the rendering and causes the synchronized output of the audio data and display of the one or more 2D images.

19. The sending computing system of claim **13**, wherein the rendering is performed by the reconstruction system and